

“つなぐチカラ”を進化させるKDDIの挑戦： AIデータセンター × GDC × Gemini

KDDI株式会社
グーグル・クラウド・ジャパン合同会社

2025/8/5

桜井 敦史

KDDI株式会社

先端技術企画本部

サービスプラットフォーム企画室長



Mao Kano

グーグル・クラウド・ジャパン合同会社

カスタマーエンジニア



01 KDDI AI DATA CENTER

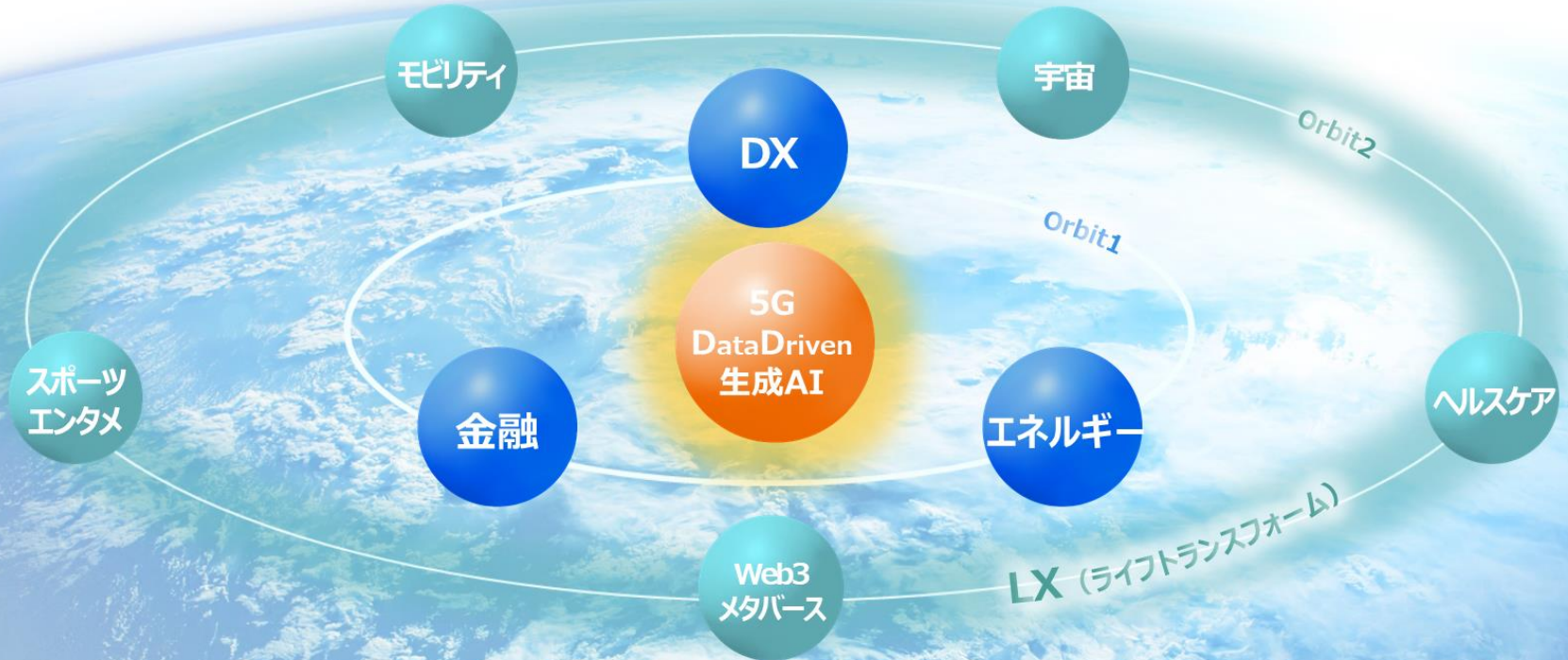
02 Gemini for GDC

03 今後のビジョン

KDDI AI DATA CENTER

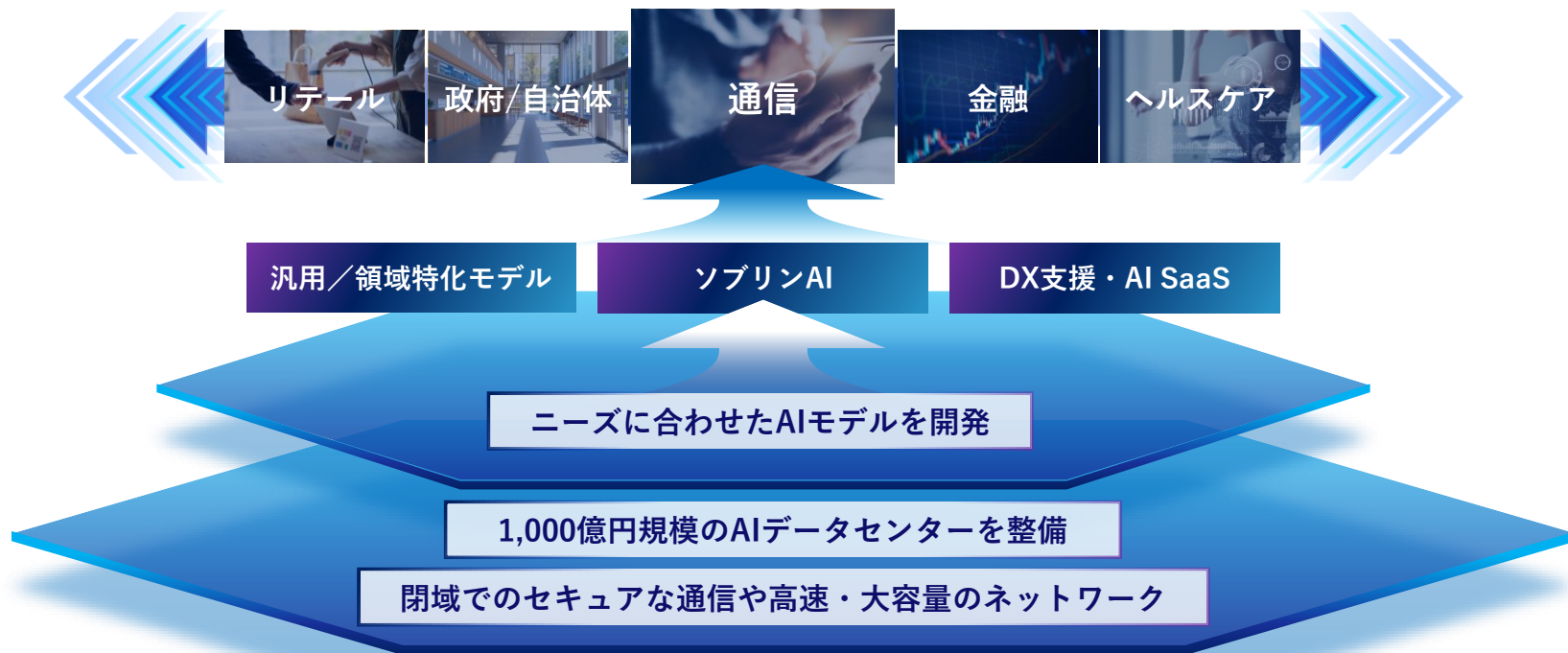
KDDIの事業戦略

5Gとデータ/生成AIを中核に、通信と周辺領域の事業成長を加速



AI事業の概要

高速・セキュアな通信インフラと大規模なAIインフラを融合
さまざまな産業や国の重要サービスを支えるプラットフォームへ



重要インフラに関する政策動向

経済安全保障と国際競争力の確保に向け、
AIをはじめとする重要インフラは他国依存から脱却する流れ

諸外国のAI関連政策

米国	インフラ投資 AI法規制緩和	民間・政府横断 AIインフラ開発による 国家安全保障確保
中国	国家体制 AI開発推進	国家戦略として AIチップ自給を強化
欧州	AI大陸行動計画 国際デジタル戦略	AIファクトリー構築 技術主権の構築
英国	Sovereign AI Unit AISI インフラ法制	AIチップ・モデル自立確保 AI安全保障インフラ整備

日本政府の動向

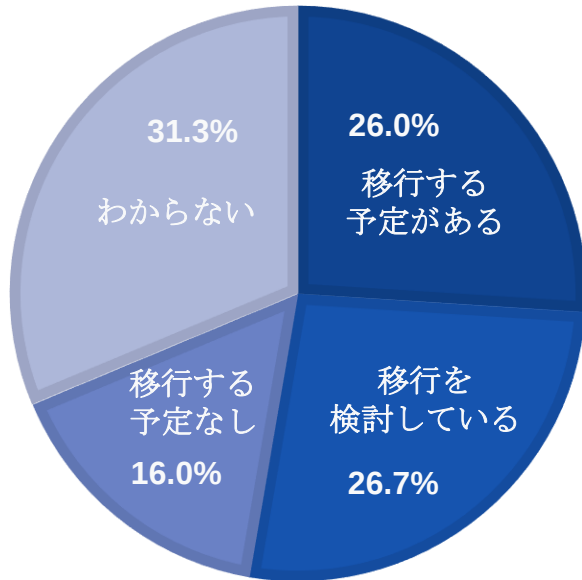
日本	官民連携 AI戦略の推進	国産AI基盤モデルの開発 (GENIAC : 2024年2月発足)
	法整備	サイバー安全保障の法制化 (能動的サイバー防衛 : 2025年5月成立)
	技術育成・投資に よる基盤強化	経済安全保障 アクションプラン (2025年5月改定)

自国で制御可能な体制＝ソブリン性の確保が
重要テーマ

クラウド普及の先にあるオンプレの価値

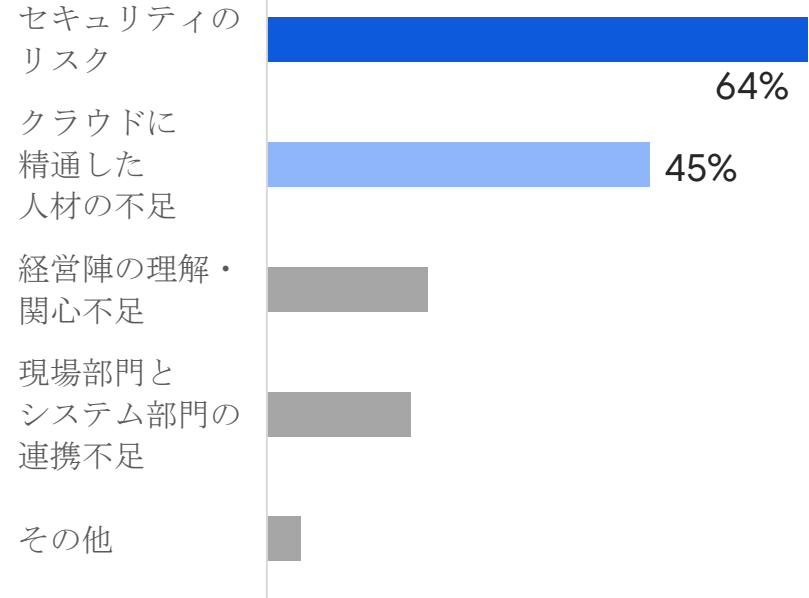
IT環境をクラウドからオンプレへ移行する流れも存在
機微なデータはクラウド利用（≒AI活用）の障壁であり、バランスがポイント

一部のIT環境でオンプレ回帰



過半数に
オンプレ回帰
の意向

クラウド導入・利用の懸念事項



KDDI AIデータセンター

25年4月に旧液晶工場を取得、AIデータセンターとして再構築
Google Cloud と戦略的提携を締結し、KDDIのDC内にGemini for GDC を展開

2025年4月シャープ堺工場
売買契約締結



Google Cloud  Gemini

GPUと水冷の データセンター



MWC2025@バルセロナで展示

スピード重視のAIデータセンター戦略

従来**3-4**年かかる工期を着工から1年以内に完遂
大阪堺データセンターとして**2025**年度内に竣工予定



GPUaaS | サービス概要

エクサスケール・コンピューティングを初期費用無しで提供
KDDIの大容量・セキュアなネットワークを自由に選択可能

GPUスーパーコンピュータ



NVIDIA H100対比

LLM推論 ▶ **30倍**

LLM学習 ▶ **4倍**

電力効率 ▶ **25倍**

NVIDIA GB200 NVL72 by HPE

1ラックで180PFLOPSの計算力※

※FP16 疎性なし

高速ストレージ



NVIDIA RA

GPUDirect Storage(GDS)対応

大容量ネットワーク



Internet
1Gbps, 10Gbps



閉域接続
WVS2



Multi Cloud接続
GCP, AWS,
Azure, etc

ソブリン性と利便性を追求

データ主権に対応したソブリンクラウドを**25年7月**にリリース
重要サービスに求められる主権レベルにあわせたサービスへ進化させる

KDDI 暗号鍵管理サービス
for Google Cloud

Gemini for GDC

GPUaaS

KDDI iret

KDDI Google Cloud

KDDI

'25/7

データ主権



データの場所・アクセス・暗号化を
自国内で管理する

自社のデータのアクセスや暗号化
について責任を有する

Coming
Soon

運用主権



クラウドの運用者を
自国内で管理する

プロバイダー側の運用に関する
透明性やコントロールを有する

Coming
Soon

技術主権



クラウドのソフトウェアを
自国内で管理する

プロバイダーへ依存せず
ソフトウェアを利用できる

Gemini for GDC

オンプレミスの生成AIの必要性

企業は**オンプレミス**で
生成**AI**を必要とするが、

利用できるモデルの
選択肢が限られる



データの保存/処理をオンプレミスで行う必要がある

クラウドに移動できないデータを扱うため、
オンプレミスで生成AIの展開が必要



モデルの希少性

高パフォーマンスのプロプライエタリモデルはオンプレミスで利用不可
OSSモデル(例:Llama/Mistral)はデプロイと管理に工数がかかる



AIリソース（HW/SWなど）の専門知識不足

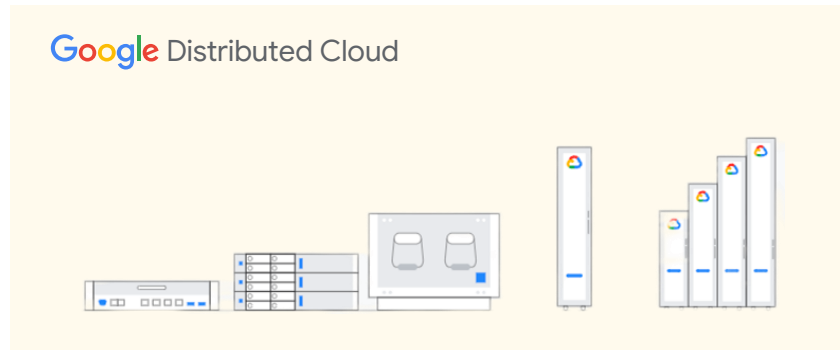
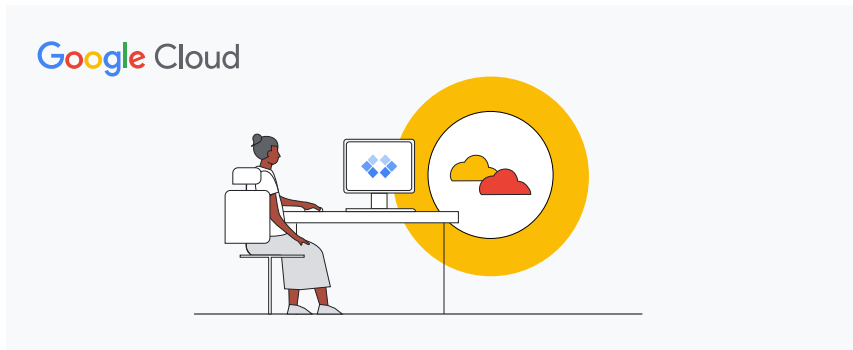
アプリケーションを内製、維持する際に影響を及ぼす



Gemini for GDC

GoogleとGoogle Cloudの高いAIケイパビリティを
お客さまのデータセンターに提供可能

GCPとGDCの違い



データの保存/処理場所が**パブリッククラウド**

データの保存/処理場所が**オンプレミス**



Vertex Model Gardenにより、豊富な**Google 1P, 3P, OSSモデル**を提供

(※GDCの提供モデルは順次拡張予定)



Google が管理するインフラや**Vertex AI** により、**最先端のAIサービス**をより簡単に利用可能

あああ

(※GDCの提供サービスの詳細は公式ドキュメント参照)



Google Cloudの標準APIとUser Experienceにより、**開発スピードと運用効率を最適化**

生成AIプラットフォームを必要な場所に提供

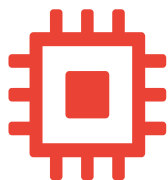


データ セキュリティ

データをオンプレミスで
保管/処理するため、
規制要件を満たします



Cloud Storage



強力な コンピューティング

最先端な
AI Hyper Computer構成を
サポート



GPUs



モデルの ライフサイクル管理

オープンで包括的なAIスタックで、
モデルのライフサイクル
(MLOps/グラウンディング/RAI)
をサポート



TensorFlow



PyTorch



最高品質な AIモデル

Google最新の1P Model,
Partner/OSSの
3rd Party Modelを
幅広くサポート



エージェント Agent

Google/Partnerによる、
エージェントの構築と
デプロイするための
包括的なツールを提供

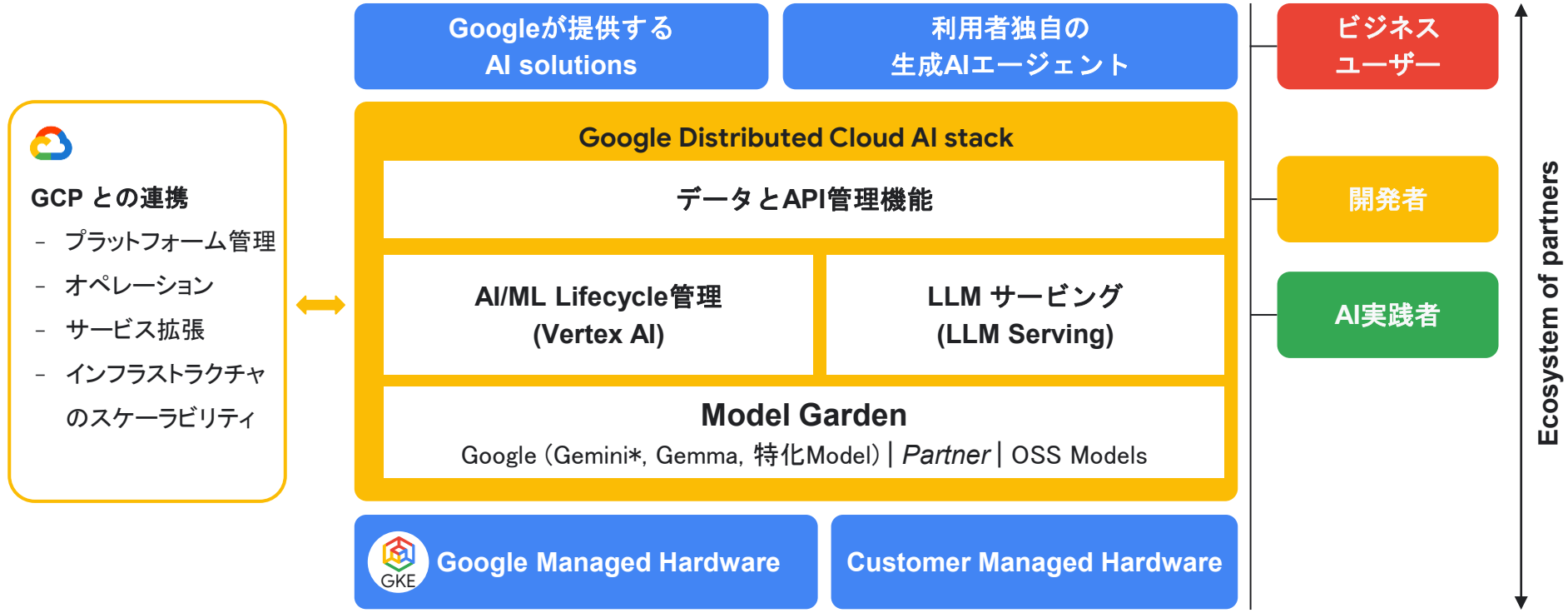


Vertex AI



Google Agentspace

GDCのAI Stack



Gemini for GDC

サービスの特徴



Googleの最高のGeminiモデル



長いコンテキストウィンドウ



マルチモーダルと多言語サポート

主なユースケース



マルチモーダルの
入力情報を理解する



GenAI検索（RAG）と
マルチターンQ&A



長いコンテキストで
バックグラウンドを理解



規制環境下の
Agentic AIの実現



既存コードの分析と
コードの生成



ドキュメント翻訳

今後のビジョン

両社で創造する価値

機微データや重要サービスにAIを活用したい企業・国の機関に
ソブリン性と利便性を両立した新しい選択肢を提供



KDDIが管理するDCで
データの保存・処理可能



Googleのサービスを
オンプレ環境で提供

OSSモデル
汎用/領域特化モデルを提供



Googleの
Geminiモデルを提供

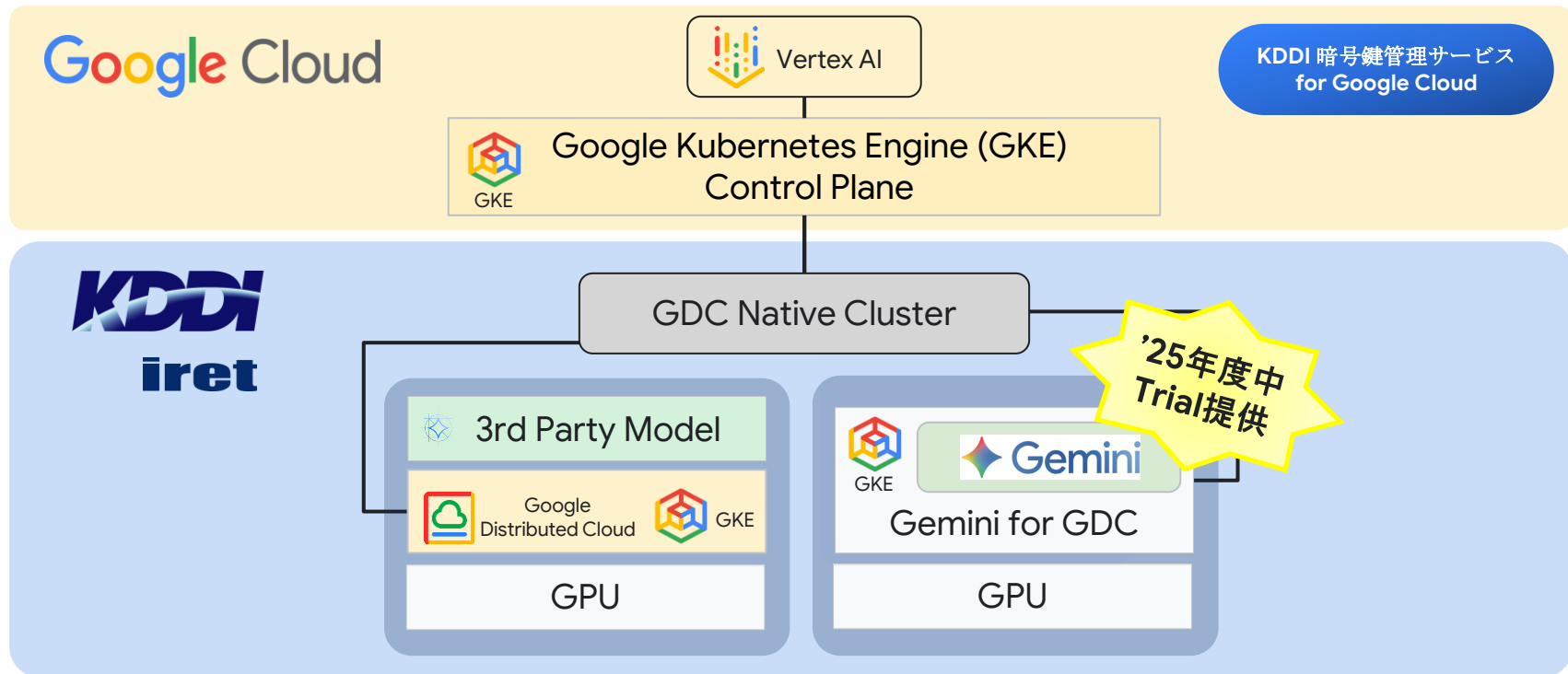
プラットフォーム導入負荷と
保守コストを軽減可能



Googleマネージドの
AIプラットフォームを提供

ソブリンAIのプラットフォーム

KDDIが管理するGPU上にGeminiやGKE環境を展開
Gemini for GDCは '25年度下期にトライアル環境を提供



ユースケース | ロボット動作の自動化

Geminiを使った**AI**ロボットがコンビニの品出しや陳列を自動化
店舗カメラの映像をオンプレ環境で管理しながら、**AI**の活用が可能



ユースケース | 通信インフラの自動化

AIエージェントが機密性の高い通信設備のデータを使って作業を自動化

Step 2: xr-2でのeBGP設定

```
configure
router bgp 65002
  bgp router-id 10.2.1.3
  address-family ipv4 unicast
  !
  neighbor 10.2.1.2
  remote-as 65001
  address-family ipv4 unicast
  route-policy PASS in
  route-policy PASS out
  !
commit
```

3. 確認手順

BGPネイバー状態確認

```
show bgp ipv4 unicast summary
show bgp neighbors 10.2.1.x
```

BGPテーブル確認

```
show bgp ipv4 unicast
show route bgp
```

接続性確認

```
ping 10.2.1.x
```

4. 期待される結果

- BGPネイバーがEstablished状態になる
- BGPテーブルに経路が表示される
- 相互通信が可能になる

BGP接続状況:

- xr-1 BGP router id 10.2.1.2, AS 65001

接続性確認結果:

```
ping 10.2.1.x
P: 100% (1/1) 0.00ms
```

設定結果サマリー:

- 手順書作成: eBGP設定手順書を詳細に作成
- 設定実行:
 - xr-1 (AS 65001) ← xr-2 (AS 65002)
 - 接続セグメント: 10.2.1.0/24
- 動作確認:
 - BGPネイバーがEstablished状態
 - 相互ping通過OK
 - BGP設定が正常に動作

ステータス:

- BGP接続設定完了
- 手順書作成から実行まで成功しました!

右側ターミナル出力:

```
RP/0/0/R0/CPU0:xr-1#sh bgp ipv4 uni summary
Mon Jul 21 04:36:23.121 UTC
% BGP Instance 'default' not active
RP/0/0/R0/CPU0:xr-1#sh run router bgp
Mon Jul 21 04:36:22.865 UTC
% No such configuration item(s)

RP/0/0/R0/CPU0:xr-1#sh RP/0/0/R0/CPU0: Jul 21 04:37:41.918 UTC: ssh cycling proxy(1100): XSECUR
ITY-5940 SYSLOG 800-6-INFO: GENERAL : ssh(940): Accepted authentication for cisco from
172.30.0.1 port 4054 ssh2
RP/0/0/R0/CPU0: Jul 21 04:37:42.650 UTC: config(60001): %BGP-CONF-5-DB_COMMIT : Configu
ration committed by user 'cisco'. Use 'show configuration commit changes 1000000016' to
view the changes.
RP/0/0/R0/CPU0: Jul 21 04:37:43.882 UTC: bgp(1000): %ROUTING-BGP-5-ASMVC_IPC_STATUS : bpe-
default:(A)Init-Id 0, Connection Open
RP/0/0/R0/CPU0: Jul 21 04:37:44.083 UTC: bgp(1000): %ROUTING-BGP-5-ASMVC_IPC_STATUS : defa
ult: process instance 1:(A)Init-Id 0, Connection Established
RP/0/0/R0/CPU0: Jul 21 04:37:44.882 UTC: bgp(1000): %ROUTING-BGP-5-ASMVC_IPC_STATUS : defa
ult:(A)Init-Id 0, Initial Config Done
RP/0/0/R0/CPU0: Jul 21 04:37:45.378 UTC: config(60001): %BGP-SYS-5-CONFIG_I : Configured
from console by cisco on vty0 (172.30.0.1)
RP/0/0/R0/CPU0: Jul 21 04:37:57.454 UTC: bgp(1000): %ROUTING-BGP-5-ADJCHANGE : neighbor 10
.2.1.3 up (vrf: default) (AS: 65002)
RP/0/0/R0/CPU0: Jul 21 04:37:57.415 UTC: bgp(1000): %ROUTING-BGP-5-NSR_STATE_CHANGE : Chan
ged state to Not NSR-Ready
```



モバイル通信の普及が社会を変えたように
AIが徐々に生活に馴染んでいく未来を
夢中になれるパートナーと共に実現

KDDI Tech noteで最新技術を発信中！ AIデータセンター × Gemini for GDCの取り組みも掲載予定 是非、ご覧ください！

KDDI Tech note



A screenshot of the KDDI Tech note website. The header features the KDDI logo, a search bar with the text "キーワードやクリエイターで検索", and navigation links for "ホーム", "記事", "マガジン", and "月別". The main banner reads "新しい「デジタル」で未来を拓ける". Below this, the "KDDI Tech note" section is highlighted, describing it as a platform for sharing KDDI's latest technology. A "フォロー" (Follow) button is visible. The page displays three featured articles: "最新情報をチェック！ KDDI 露出情報" (Latest information check! KDDI exposure information), "AIデータセンターの要は「冷却」にあり！ 水冷検証環境をご紹介します" (The key to AI data centers is 'cooling'! Introducing our water-cooled testing environment), and "AI基本社会を目指す韓国新政府の動き" (Movement of the new South Korean government aiming for AI basic society). Each article includes a thumbnail, title, and a "KDDI Tech note" label with a date.

ご清聴ありがとうございました

アンケートのご回答もよろしくお願いいたします